# Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

# PRACTICAL APPLICATION OF STANDARD TESTS IN SPELLING, LANGUAGE, AND ARITHMETIC

JOSEPH P. O'HERN
Assistant Superintendent of Schools, Rochester, New York

Rochester has done so much for its school children along the lines of manual training, domestic science, domestic art, music, drawing, physical education, etc., that the question has often been raised as to whether or not the subjects of reading, spelling, arithmetic, language, etc., were receiving the necessary time and attention. Monthly tests and city-wide term examinations indicated a satisfactory condition. These tests, however, were made up within the system and were therefore local. The use of standard tests was started in order to find out how Rochester would compare with other cities in the so-called formal subjects. The more we use standard tests the more we are convinced that these tests are useful, not only for the purposes of diagnosis and comparison, but also for the purpose of locating new pupils in the right grade. This is especially true of the Judd Arithmetic Test, which may be used in grading new pupils who come from other school systems at all times of the year.

## SPELLING TESTS

In spelling three tests were given: the Buckingham ten-point spelling scale and two Ayres spelling tests. In the Buckingham test the scores were high in all grades from the third to the eighth inclusive.

For the first Ayres spelling test the same words were chosen that were given in Springfield, Butte, Oakland, and Salt Lake City.

The second Ayres spelling test was the same as that chosen by the Cleveland Survey Committee. The low score in the fourth grade was due to the fact that a number of words in the examination were not taught in this grade. In general, however, most of the words were found in our course of study for the grade concerned.

The results of these standard spelling tests justified the time limit we had set for spelling, which is about 2 per cent less than the time given to spelling in most city school systems.

## TABLE I

COMPARISON OF THE RESULTS FROM THE BUCKINGHAM SPELLING SCALE IN ROCHESTER AND NEW YORK

| | GRADE | | | | | | AVERAGE |
|---|---|---|---|---|---|---|---|
| | III | IV | V | VI | VII | VIII | |
| New York...... | 30 | 47 | 60 | 70 | 78 | 85 | 62 |
| Rochester....... | 55 | 70 | 80 | 85 | 91 | 95 | 77 |
| Difference... | +25 | +23 | +20 | +15 | +13 | +10 | +15 |

## TABLE II

COMPARISON OF THE RESULTS FROM THE FIRST AYRES SPELLING TEST WITH OTHER CITIES

| | GRADE | | | | | | AVERAGE |
|---|---|---|---|---|---|---|---|
| | III | IV | V | VI | VII | VIII | |
| Springfield, Ill.......... | 65.0 | 70.0 | 72.0 | 68.0 | 73.0 | 75.0 | 70.0 |
| Butte, Mont........... | 81.8 | 78.7 | 84.5 | 75.0 | 76.2 | 89.4 | 80.3 |
| Oakland, Cal........... | 66.7 | 75.9 | 84.7 | 80.2 | 79.9 | 76.3 | 76.5 |
| Salt Lake City, Utah.... | 89.9 | 78.8 | 87.6 | 86.8 | 87.1 | 82.2 | 86.0 |
| Rochester, N.Y......... | 77.8 | 77.4 | 85.5 | 82.9 | 83.1 | 84.9 | 81.9 |

NOTE.—In Butte and Salt Lake City this test included only the B sections of the grade while in Oakland and Rochester it included both A and B sections.

## TABLE III

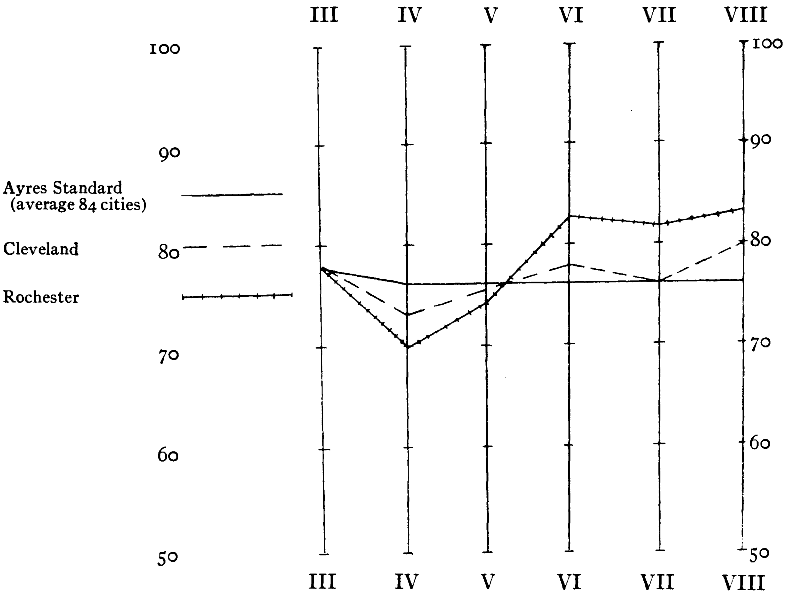
COMPARISON OF THE RESULTS FROM THE CLEVELAND SURVEY AYRES SPELLING TEST WITH THE AYRES STANDARD

| | GRADE | | | | | | | | | | | | AVERAGE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | III | | IV | | V | | VI | | VII | | VIII | | |
| | B | A | B | A | B | A | B | A | B | A | B | A | |
| Standard.......... | 73.0 | 81.4 | 73.0 | 78.9 | 73.0 | 78.9 | 73.0 | 78.9 | 73.0 | 78.9 | 73.0 | 78.9 | 76.0 |
| Rochester......... | 73.2 | 81.9 | 68.6 | 71.4 | 71.5 | 76.8 | 80.9 | 84.6 | 79.5 | 85.2 | 81.6 | 84.9 | 77.5 |
| Difference..... | +0.2 | +0.5 | −4.4 | −7.5 | −1.5 | −2.1 | +7.9 | +5.7 | +6.5 | +6.3 | +8.6 | +6.0 | +1.5 |

### TABLE IV

COMPARISON OF THE RESULTS FROM THE CLEVELAND SURVEY AYRES SPELLING TEST WITH THE ROCHESTER SCORES

| | GRADE | | | | | | AVERAGE |
|---|---|---|---|---|---|---|---|
| | III | IV | V | VI | VII | VIII | |
| Cleveland | 78.0 | 73.0 | 75.0 | 78.0 | 76.0 | 80.0 | 76.0 |
| Rochester | 77.5 | 69.9 | 74.0 | 82.5 | 81.9 | 83.3 | 77.5 |
| Difference | −0.5 | −3.1 | −1.0 | +4.5 | +5.9 | +3.3 | +1.5 |



## LANGUAGE TESTS

The Kansas Silent-Reading Test was given in all grades of the city, from the third to the eighth. The scores were slightly above the standard in the third, fourth, and fifth, slightly below in the sixth, equal in the seventh, and somewhat below in the eighth grade. The silent-reading test showed that the schools having a large foreign population fell below the standard median, while schools in the English-speaking sections ran well above the standard.

TABLE V

COMPARISON OF THE RESULTS FROM THE KANSAS SILENT-READING TEST WITH THE STANDARD

| | GRADE | | | | | |
|---|---|---|---|---|---|---|
| | III | IV | V | VI | VII | VIII |
| Standard | 5.0 | 9.4 | 13.4 | 13.8 | 16.5 | 19.2 |
| Rochester | 5.9 | 9.9 | 13.7 | 13.6 | 16.5 | 17.4 |
| Difference | +0.9 | +0.5 | +0.3 | −0.2 | 0 | −1.8 |

### THORNDIKE'S VISUAL-VOCABULARY TEST

In order to test the power of word-recognition and word-meaning the Thorndike Visual-Vocabulary Test was given in eleven schools in Grades III B to VIII A inclusive, and in the B and A divisions of the third grade in eight additional schools, all largely foreign except one.

TABLE VI

COMPARISON OF THE RESULTS FROM THE THORNDIKE VISUAL-VOCABULARY TEST WITH EIGHTEEN CITIES OF INDIANA

| | GRADE | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | III | | IV | | V | | VI | | VII | | VIII | |
| | B | A | B | A | B | A | B | A | B | A | B | A |
| Rochester | 4 | 4 | 4½ | 5½ | 5½ | 6 | 6½ | 7 | 7½ | 8½ | 8½ | 8 |
| Indiana | 4 | | 5.26 | | 6 | | 6.66 | | 7.29 | | 7.91 | |

In the visual-vocabulary test Thorndike attempts to test the understanding of the meaning of words. There are nine steps in this scale, five words each, the steps being equal in difficulty. There is increasing difficulty from line to line. The words were selected from a long list and tested at first by the examination of about fifteen hundred children. The pupils use a letter to show what the word represents, *A* for animal, *F* for flower, etc. This test is valuable because it determines whether the child grasps the meaning of the printed word. It gives undue prominence to names of animals and flowers. No standard grade scores have been determined. It is impossible, therefore, to compare the record of the Rochester schools with a standard score. A record, however, was made in eighteen cities in Indiana that may serve for the time being for such a comparison. On that basis the record in Rochester is strikingly similar to that of the cities of Indiana.

The record in the various schools is very even and consistent from grade to grade, except in the foreign schools, where thirteen groups failed to score at all in the III B and six groups failed to score in III A.

The principal in a local school, which is 90 per cent Italian, made a very complete report of the results attained in his school. The final statement in this report might be quoted as an explanation of the failure of so many groups to score in III B and III A.

> This scale was uninspiring from the standpoint of results. The element of nationality enters even to a greater extent than in the Trabue scale. Children who hear little or no English spoken at home and whose opportunities for reading outside of school are very limited had a hard time with it. Such children's vocabulary of words taken apart from the context of a paragraph is very limited.
>
> Some erratic tendencies are noticed in these results. A great many individual children succeeded in passing certain steps after having failed in one, two, or three of the preceding steps. To some extent this may have been due to the type of words occurring in the steps. It was very evident from the results that words belonging to the general classes of animals or flowers were much more familiar than the words expressing time or the moral qualities classified under "Good" or "Bad."
>
> The fact that the IV A results were better uniformly than those from the V B, and the V A results better than all others, reveals a condition requiring further study. These results are in accord, however, with those attained by the corresponding grades in the Trabue test. In general the over-age children did poorly in this test. This was to be expected, however, since in a majority of cases it is inability to use English that has caused their retardation.
>
> At least two lines of remedial effort are suggested by these results, viz., (1) more emphasis on the meaning of words occurring in reading and spelling lessons with practice in using the same in sentences; (2) more dictionary work in intermediate and upper grades.

### THORNDIKE'S SCALE ALPHA

The Thorndike Scale Alpha consists of a series of paragraphs. The child's understanding is tested by requiring him to answer questions and to follow directions concerning the paragraph. The scale is designed for pupils from the third grade to the high school. It lacks the advantage of a connected story, and the suitability of subject-matter might possibly be questioned as being devoid of interest to many pupils.

As there was some question as to the method of scoring the results in this test, it was decided for the sake of uniformity that the total class average of correct answers to questions would be used as the score.

Since no standard scores are available the results are chiefly valuable for purposes of internal comparison. In the schools where the test was given in all grades the record for the number of divisions of a grade which were above or below the median is shown in Table VII. Schools C, D, G, and K in this table are predominantly foreign in nationality, Schools E, H, and I are mixed in nationality, Schools A, B, and F are American, and School J is a village school recently annexed to the city. This comparison would seem to prove the point quoted above, that the element of nationality is a very strong factor. The schools having the larger number of divisions above the median might be considered as favored communities in point of environment, home influences, and opportunities.

TABLE VII

| School | Above Median | Below Median | School | Above Median | Below Median |
|---|---|---|---|---|---|
| A | 11 | 1 | G | 1 | 7 |
| B | 11 | 1 | H | 8 | 4 |
| C | 4 | 8 | I | 6 | 2 |
| D | 1 | 7 | J | 0 | 12 |
| E | 6 | 6 | K | 0 | 4 |
| F | 10 | 2 | | | |

In comparing the results of the visual-vocabulary test with the results of the scale-alpha test it is evident that word-recognition is easier than content-reproduction. It may be that too many of our school children learn to pronounce words without being made fully conscious of their meaning, but in the absence of standards it is a problem for further investigation.

The results obtained from the use of scale alpha would seem to emphasize the fact that silent reading with understanding is more important and should receive more attention than oral reading.

TABLE VIII

THORNDIKE SCALE ALPHA

| | GRADE | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | III | | IV | | V | | VI | | VII | | VIII | |
| | B | A | B | A | B | A | B | A | B | A | B | A |
| Rochester median.. | 8.3 | 12.1 | 13.9 | 15.8 | 17.4 | 17.95 | 18.7 | 18.9 | 13.9 | 15.4 | 16.4 | 17.6 |

## TRABUE COMPLETION LANGUAGE SCALE B

The Trabue Completion Language Scale B was given in eleven schools from Grades III B to VIII A inclusive and in the B and A divisions of the third grade in the seven schools that comprise the Washington Junior High School District. The eleven schools represent all the types of schools to be found in a city the size of Rochester, the foreign in nationality, the mixed in nationality, and the well-to-do sections.

The Trabue test is merely a list of incomplete sentences arranged in such a way that pupils may be expected to complete them, hence the name. There are ten sentences in each test. Seven minutes are allowed for each test. Two credits are allowed for each sentence. These tests are considered of value in judging general intelligence.

Dr. Trabue found that the results of his tests agreed with the opinion of the teacher as to the pupils' general ability.

Of these completion-test language scales Dr. Trabue says:

> No attempt has been made to define language in any strict sense, and it is entirely possible that some persons may be able to speak the English language and perhaps to write it fairly well without being able to make a very high score on these scales. It may also happen that some individuals will be found who score well on these language scales and are yet unable to quote a

single rule of English grammar. On the whole, however, it will be found that ability to complete these sentences successfully is very closely related to what is usually called language ability.

The results of the test in the eleven schools were eminently satisfactory from the point of view of achievement (see Table IX). Only four groups out of the one hundred and forty groups examined were below the standard median. The gain in score from lower to higher grades was uniformly consistent. This would tend to confirm the assumption of the author, in arranging the language scales, "that older children will not only be able to do the same tasks more rapidly and more perfectly than younger children, but that the older children will also be able to accomplish perfectly tasks which the younger children could not begin to do."

In only seven cases did a B group of a grade within a school score a higher standing than the A group of the same grade. While there is variation in the scores within a grade, school by school, still the variation is not so marked as to point to lack of uniformity.

In this test the foreign schools in the upper grades vary little from other schools, and in some groups attain a higher median.

TABLE IX

COMPARISON OF THE RESULTS FROM THE TRABUE COMPLETION TEST WITH THE STANDARD

| | GRADE | | | | | |
|---|---|---|---|---|---|---|
| | III | IV | V | VI | VII | VIII |
| Rochester | 7.77 | 10.17 | 11.55 | 12.57 | 14.37 | 14.75 |
| Standard | 6.00 | 8.00 | 9.80 | 11.00 | 12.30 | 13.30 |
| Difference | +1.77 | +2.17 | +1.75 | +1.57 | +2.07 | +1.45 |

## STARCH GRAMMATICAL SCALE A

The Starch Grammatical Scale A was given in the seventh and eighth grades in eight schools. The scale includes different items of grammatical forms and is therefore only a general measure of ability to use correct language forms. Each pupil is given a printed copy of the scale and is allowed as much time as he needs. The scale directions are as follows: "Each of the following sentences gives in parenthesis two ways in which it may be stated. Cross out the one you think is incorrect or bad. If you think both are incorrect, cross both out. If you think both are correct, underline both."

The results point to a condition that is unusual, especially in the seventh grade, where the score was equal to the score of the eighth grade. In fact, the VII A score was higher than the score of VIII B and VIII A.

The difference between VII A and VIII A scores would seem to indicate too little emphasis in VIII A and perhaps a corresponding overemphasis in VII A. The question of nationality or neighborhood does not seem to enter. It is a question of balance and right emphasis.

TABLE X

COMPARISON OF THE RESULTS IN THE STARCH GRAMMATICAL SCALE WITH THE STANDARD

| | GRADE | | | |
|---|---|---|---|---|
| | VII | | VIII | |
| | B | A | B | A |
| Rochester........ | 8.4 | 8.9 | 8.45 | 8.8 |
| Standard......... | 8.0 | | 8.3 | |
| Difference.... | +.6 | | +.3 | |

## JUDD-COURTIS ARITHMETIC TEST

The course of study affirms "on the one hand that the study of arithmetic should develop mental power along the lines of attention, concentration, judgment, and reasoning; on the other that it should give the child such skill in the use of numbers as will enable him to perform the fundamental operations with facility and accuracy and in such a manner that the mere mechanics of these processes shall not retard his thinking." The need of adequate ways and means of testing instruction in arithmetic has been felt for many years. The pioneers in the field of testing in the fundamental operations were Rice and Stone. Courtis came next with his Series A, which, however, proved unsatisfactory and was replaced by Series B, which represented a distinct improvement over the earlier form. Series B has been used extensively throughout the country, and as a result reliable standards have been established in the attainments of children in the fundamental operations of arithmetic. It was felt by the Cleveland Survey Committee that "it was necessary to introduce, between the very simple type of example in the first series (Courtis Series A) and the highly complex type in the second series (Courtis Series B), tests representing types of intermediate complexity." As a result, and with the co-operation of Mr. Courtis, the Judd arithmetic tests were devised and given to the B grades in Cleveland and later to the B and A grades in Grand Rapids. The Judd test is composed of four sets in addition (A, E, J, M), two in subtraction (B, F), three in multiplication (C, G, L), four in division (D, I, K, N), and two in fractions (H, O). The pupil begins with Set A and takes each set in order. The test is spiral in character, and in this respect it differs from the Courtis tests.

> The several sets in each operation are arranged in the test in the order of their complexity, but with them are interwoven the sets of the other operations. Thus a pupil first works on a set of examples in addition, then passes successively to sets in subtraction, multiplication, and division before encountering addition again. This changing from one type of operation to another lessens the strain on the pupil which is involved in a prolonged test of this sort.[1]

The Judd test was chosen by the Cleveland Survey Committee because it was thought that "it would, on the one hand, show the general standing of the city as a whole in the fundamentals of arithmetic and would, on the other hand, be diagnostic in its character, indicating school, class, and individual weakness in each of the different types of operations which enter into the solving of the more complex examples in each of the four fundamental operations."

The Judd test was given in the same schools as the other standard tests. In Cleveland the test was given to the B grades only and in Grand Rapids to the B and A grades. Rochester medians are below the Cleveland medians in every B grade in Sets B, H, and K, but above in every grade in Sets L and M.

---

[1] George S. Counts, *Arithmetic Tests and Studies in the Psychology of Arithmetic* (Supplementary Educational Monographs, Vol. 1, No. 4, whole No. 4), p. 6.

In B grades 27 medians were above Cleveland medians
“ B “ 50 “ “ below “ “
“ A “ 49 “ “ above “ “
“ A “ 35 “ “ below “ “

Medians improve from grade to grade in every division of the test. Occasionally the record of the A grade is less than the corresponding B grade, but only slightly less in any case.

Compared with Grand Rapids—

In B grades 36 medians were above Grand Rapids medians
“ B “ 29 “ “ below “ “ “
“ B “ 3 “ “ equal to “ “ “
“ A “ 44 “ “ above “ “ “
“ A “ 26 “ “ below “ “ “
“ A “ 1 median was equal to “ “ “

The comparison with Grand Rapids is more favorable to Rochester than the comparison with Cleveland. This may be due to the fact that there is more certainty as to the reliability of the Grand Rapids scores. Professor Counts says: “A great deal more care was taken here than in Cleveland to insure the results against error. In Cleveland the teachers were inexperienced in giving tests, while in Grand Rapids they were all more or less familiar with the Courtis tests.” In Rochester the tests were given and timed by principals who had taken the courses in tests and measurements under Professor L. A. Pechstein, at the University of Rochester. Testing conditions were standardized and carefully supervised.

A more important way of looking at the scores is from the point of view of difficulty. In point of difficulty from simple to complex Professor Counts groups the sets as follows:

1. A, B, C, D
2. E, F, G
3. H
4. I, J, K
5. L, M, N, O

Compared with Cleveland in—

| | | | | | |
|---|---|---|---|---|---|
| 1. A, B, C, D | 4 | medians | above | Cleveland | medians |
| | 20 | “ | below | “ | “ |
| 2. E, F, G | 4 | “ | above | “ | “ |
| | 10 | “ | below | “ | “ |
| 3. H | 0 | “ | above | “ | “ |
| | 4 | “ | below | “ | “ |
| 4. I, J, K | 7 | “ | above | “ | “ |
| | 8 | “ | below | “ | : |
| 5. L, M, N, O | 12 | “ | above | “ | “ |
| | 7 | “ | below | “ | “ |

Compared with Grand Rapids in—

| | | | | | | |
|---|---|---|---|---|---|---|
| 1. A, B, C, D | 25 | medians | above | Grand | Rapids | medians |
| | 17 | “ | below | “ | “ | “ |
| | 2 | “ | equal | “ | “ | “ |
| 2. E, F, G | 19 | “ | above | “ | “ | “ |
| | 11 | “ | below | “ | “ | “ |
| 3. H | 0 | “ | above | “ | “ | “ |
| | 7 | “ | below | “ | “ | “ |
| 4. I, J, K | 18 | “ | above | “ | “ | “ |
| | 9 | “ | below | “ | “ | “ |
| 5. L, M, N, O | 18 | “ | above | “ | “ | “ |
| | 11 | “ | below | “ | “ | “ |
| | 2 | “ | equal | “ | “ | “ |

The foregoing comparisons would seem to indicate that in the more difficult sets of the test Rochester pupils have a slight advantage. Whether a larger expenditure of time on the fundamental operations in the lower grades would seem necessary from the results of this one test is a question. The deciding factor should be how well the work is done rather than how much time is taken. In all probability too much time is taken in some cases and too little in other cases. Some pupils and some groups need more time and other pupils and other groups less time. Of this the teacher must be the judge, save where a test clearly proves the direct necessity of a definite course of action.

A comparison by grades and tests is given in Table XI (see p. 677). From this table the following details may be selected for comment:

### SETS A, E, J, M—ADDITION

*Set A*.—The medians are below Cleveland in all but IV B and V B. The medians are above Grand Rapids in III B and A, IV B and A, V B and A, VI B, equal in VI A, and below in VII B and A and VIII B and A, indicating a need of daily rapid drill in the upper grades in order to maintain the skill developed in the lower grades.

*Set E*.—This is short, single-column addition. It is evident that insufficient attention is given to this type of written work.

*Set J*.—This is long, single-column addition. The Rochester medians are above Cleveland in all grades except IV B and V B, and above Grand Rapids in five and below in four divisions. If the same attention had been given to short-column written addition as to long single-column addition the test would have shown the same satisfactory results.

*Set M*.—Above Cleveland in every group and below Grand Rapids in three divisions, V A, VII B, and VIII B. This shows ample practice in written work.

### SETS B, F—SUBTRACTION

*Set B*.—Many pupils added Set B instead of subtracting, which may explain the poor record in this set.

*Set F*.—This does not mean that children cannot have examples under the Cleveland form. (See *Course of Study*, p. 14, sec. 3D.)

### SETS C, G, L—MULTIPLICATION

*Set C*.—Below Cleveland in B grades and above in A grades, above Grand Rapids in five and below in five. Pupils misunderstood this form of multiplication, as it is not used in the schools. This suggests that it might be well in grades above III B to express the work in both horizontal and vertical forms.

*Set G*.—Above Cleveland in IV B. Other grades practically same as Cleveland. Above Grand Rapids in all but three divisions.

*Set L*.—This is the most difficult set and our medians are above Cleveland in all grades B and A and above Grand Rapids in five and below in three, VII B, VIII B, and VIII A.

### SETS D, I, K, N—DIVISION

*Set D.*—Below Cleveland in IV B, VII B, and VIII B and above in V B and VI B, and above Grand Rapids in all but one. This may indicate that upper grades have not been using this form of division.

*Set I.*—Above Cleveland in IV B and V B and the same in VI B and VII B. Above in Grand Rapids in all but one. The slight falling off in VIII B might indicate a need of drill in this type.

*Set K.*—Below Cleveland in all grades, and below Grand Rapids in four and above in four. This indicates a distinct need of continuous drill on simple long division.

*Set N.*—Below Cleveland in all B grades except VI B and above in all A grades. Above Grand Rapids in all but one case, where grades are equal. This may mean the necessity of more drill on simple forms of long division.

### SETS H, O—FRACTIONS

*Set H.*—Below Cleveland and Grand Rapids in all grades B and A. It is evident that too much was taken for granted in the work in fractions. The results indicate the need of more stress on the simple forms.

*Set O.*—Below Cleveland and below Grand Rapids. Here again the need of more daily drill on simple fractions is apparent.

Principals and teachers should make a careful study of the scores, not only the scores in Rochester, but also the scores in Cleveland and Grand Rapids. How do we compare with Cleveland and Grand Rapids as a system, grade by grade and set by set? How do our own schools compare with one another? How do grades compare in the various operations, set by set? Is the progress of a grade in any school normal, or does it show variation—above the median in one operation and below in another, showing lack of balance, lack of stress, and lack of supervision? Does the record of an individual pupil show variation to such an extent that remedial measures should be sought to eradicate the weakness or to remove wrong associations in certain phases of number work? We have traced the record grade by grade and set by set and noted the cases where a grade is above in all sets, below in all sets, or above in some and below in others. In general there is a consistent and stable balance, but exceptions occur here and there. The most noticeable variation is in Set H, fractions. In V B, where Set H is first given, the variation is from 0.5 lowest to 10.1 highest; in V A from 0 to 7; in VI B from 0 to 9; in VI A from 1 to 6.8; in VII B from 0 to 6.5; in VII A from 2 to 8; in VIII B from 0 to 6.1, and in VIII A from 2.3 to 7.5. In no other set is there such pronounced variation in scores. Variations are bound to appear, owing to individual differences, teaching conditions, environment, etc., but given average intelligence and normal conditions there should not be such pronounced variation as in the case of Set H, where Rochester is below Cleveland and Grand Rapids in every grade. In Set O we are also below Cleveland and Grand Rapids in all but VIII A; still there is not the same pronounced variation grade by grade as in the case of Set H.

Attention is called to chapters v and vi in Counts's *Arithmetic Tests and Studies in the Psychology of Arithmetic*. Chapter v contains a "comparison of the arithmetical abilities of certain age and promotion groups" and chapter vi contains a "comparison of the arithmetical abilities of certain race groups." A division of a number of groups into under-age, normal, and over-age groups proved to our satisfaction that the conclusions reached by Professor Counts in chapter v are applicable to Rochester, namely, that "'fast' corresponds to the young and the 'slow' to the old." Furthermore, a glance at the arithmetic scores of the various schools will show that his conclusions as to the abilities of certain race groups would be found true in Rochester. The scores attained by the various grades in one school, 90 per cent Italian, are invariably above the median and in several grades the very highest scores of all. In V A this school attained the highest score in all but two sets. Other schools, mostly foreign, might also be mentioned in this connection. There is apparently no handicap from language or race in the study of arithmetic. This is quite natural, as number is a universal concept apart from any medium of expression save the symbol itself. As Professor Counts suggests, there may be social factors at work tending to bring about these differences. Whatever may be the explanation, the fact remains that in

## TABLE XI

### COMPARISON OF THE RESULTS OF THE JUDD ARITHMETIC TEST WITH THE CLEVELAND AND GRAND RAPIDS SCORES

| GRADE | CITY | SET | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
| IIIB ... | Cleveland.... | 13.4 | 9.3 | 6.5 | 6.3 | 4.3 | 2.0 | 2.0 | ...... | 0.6 | 1.9 | ...... | ...... | 1.4 | ...... | ...... |
| | Grand Rapids. | 11.8 | 6.3 | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| | Rochester.... | 12.3 | 4.5 | 5.0 | ...... | 4.1 | 1.9 | ...... | ...... | ...... | 1.9 | ...... | ...... | 1.8 | ...... | ...... |
| IIIA ... | Grand Rapids. | 13.4 | 8.4 | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... | ...... |
| | Rochester.... | 13.95 | 3.95 | 8.2 | 6.25 | 3.65 | 2.35 | 2.9 | ...... | 0.2 | 2.6 | ...... | 1.6 | 2.05 | ...... | ...... |
| IV B ... | Cleveland.... | 17.8 | 13.4 | 12.0 | 12.4 | 5.3 | 4.9 | 3.9 | ...... | 1.1 | 3.2 | 4.0 | 1.7 | 2.5 | 0.8 | ...... |
| | Grand Rapids. | 13.6 | 9.1 | 7.1 | 6.9 | 4.1 | 2.8 | 2.2 | ...... | 0.7 | ...... | ...... | ...... | ...... | ...... | ...... |
| | Rochester.... | 18.3 | 7.5 | 11.0 | 10.8 | 4.6 | 4.0 | 4.0 | ...... | 1.3 | 2.9 | ...... | 2.0 | 2.9 | ...... | ...... |
| IV A ... | Grand Rapids. | 16.4 | 12.1 | 11.3 | 10.4 | 4.6 | 4.1 | 3.3 | ...... | 0.9 | 2.8 | ...... | ...... | 2.3 | ...... | ...... |
| | Rochester.... | 17.5 | 13.0 | 13.2 | 11.7 | 5.6 | 6.3 | 5.0 | ...... | 1.7 | 3.6 | 2.7 | 3.0 | 3.6 | 1.0 | ...... |
| V B .... | Cleveland.... | 22.2 | 17.2 | 15.5 | 15.7 | 6.3 | 6.7 | 5.2 | 5.0 | 2.0 | 4.0 | 6.8 | 2.5 | 3.2 | 1.3 | ...... |
| | Grand Rapids. | 20.3 | 14.7 | 13.7 | 12.5 | 5.2 | 6.0 | 4.0 | ...... | 1.3 | 3.4 | 3.0 | 2.3 | 3.0 | 0.7 | ...... |
| | Rochester.... | 23.0 | 16.5 | 14.7 | 17.0 | 6.0 | 7.1 | 5.1 | 2.7 | 2.8 | 4.3 | 3.7 | 3.1 | 3.6 | 1.0 | ...... |
| V A .... | Grand Rapids. | 21.5 | 15.9 | 14.0 | 14.3 | 5.4 | 6.5 | 4.9 | 6.3 | 1.4 | 3.7 | 4.3 | 2.9 | 3.6 | 0.8 | ...... |
| | Rochester.... | 23.0 | 17.15 | 14.25 | 15.85 | 6.1 | 6.75 | 5.15 | 2.15 | 2.7 | 4.2 | 4.95 | 2.95 | 3.25 | 1.5 | 3.0 |
| VIB ... | Cleveland.... | 24.8 | 19.8 | 16.6 | 18.5 | 6.8 | 7.5 | 5.5 | 5.5 | 3.1 | 4.4 | 8.5 | 2.8 | 3.8 | 1.7 | 3.1 |
| | Grand Rapids. | 22.8 | 16.8 | 15.5 | 15.5 | 6.0 | 7.1 | 5.3 | 6.2 | 2.3 | 4.1 | 5.4 | 3.3 | 4.3 | 1.1 | 3.5 |
| | Rochester.... | 23.5 | 19.3 | 15.0 | 18.8 | 6.5 | 7.6 | 5.4 | 1.8 | 3.0 | 4.7 | 5.8 | 3.8 | 4.3 | 1.8 | 3.0 |
| VIA ... | Grand Rapids. | 25.0 | 19.1 | 17.0 | 16.9 | 6.6 | 8.0 | 5.6 | 6.5 | 3.0 | 4.5 | 6.5 | 3.6 | 4.5 | 1.4 | 3.6 |
| | Rochester.... | 25.0 | 19.55 | 16.0 | 18.75 | 6.55 | 8.5 | 5.75 | 2.55 | 3.5 | 4.9 | 7.0 | 3.8 | 4.25 | 1.85 | 3.35 |
| VIIB .. | Cleveland.... | 26.7 | 21.5 | 17.7 | 20.8 | 7.5 | 8.6 | 5.9 | 7.7 | 4.0 | 4.9 | 10.1 | 3.2 | 4.4 | 2.0 | 4.1 |
| | Grand Rapids. | 26.5 | 21.3 | 17.7 | 18.4 | 7.2 | 9.3 | 6.1 | 9.0 | 3.8 | 5.4 | 7.5 | 4.3 | 4.9 | 1.7 | 3.9 |
| | Rochester.... | 25.6 | 20.0 | 15.9 | 18.4 | 6.8 | 7.8 | 5.8 | 3.7 | 4.0 | 4.9 | 6.8 | 4.0 | 4.8 | 1.7 | 3.5 |
| VIIA .. | Grand Rapids. | 27.3 | 20.7 | 18.8 | 19.7 | 7.2 | 9.6 | 6.1 | 7.8 | 4.1 | 5.3 | 8.8 | 4.5 | 5.0 | 1.8 | 4.6 |
| | Rochester.... | 26.0 | 19.75 | 17.25 | 19.45 | 6.95 | 8.55 | 6.25 | 4.1 | 4.15 | 4.95 | 8.0 | 4.65 | 5.1 | 2.45 | 3.95 |
| VIIIB .. | Cleveland.... | 27.5 | 26.0 | 19.0 | 22.5 | 7.8 | 10.1 | 6.6 | 8.5 | 4.7 | 5.7 | 12.5 | 3.9 | 5.1 | 2.6 | 5.5 |
| | Grand Rapids. | 29.5 | 22.8 | 49.3 | 20.5 | 7.8 | 10.3 | 6.7 | 8.6 | 4.0 | 5.7 | 9.7 | 4.9 | 5.7 | 2.0 | 5.5 |
| | Rochester.... | 24.3 | 21.15 | 17.15 | 21.4 | 8.35 | 8.55 | 6.45 | 4.8 | 4.25 | 5.65 | 8.65 | 4.7 | 5.55 | 2.5 | 4.4 |
| VIIIA.. | Grand Rapids. | 30.3 | 25.5 | 20.7 | 23.0 | 8.1 | 11.0 | 6.8 | 8.8 | 4.7 | 6.5 | 10.3 | 4.9 | 5.7 | 2.3 | 4.8 |
| | Rochester.... | 29.25 | 24.8 | 17.8 | 23.15 | 7.75 | 9.7 | 7.0 | 4.9 | 5.1 | 5.8 | 9.9 | 4.8 | 6.2 | 2.9 | 5.0 |

arithmetic the pupils of foreign parentage are in the lead. Finally, it is significant to note that the types of errors found by Counts in his analysis of Cleveland and Grand Rapids scores are likewise revealed in the Rochester results.

## SUMMARY AND CONCLUSION

From this attempt to measure scientifically the educational attainment of school children in the language and arithmetical abilities eight points of major significance may be drawn in summary:

*Language abilities.*—

1. The development of language ability is being adequately secured, so far as the entire city system is concerned. This is shown consistently by all the tests used.
2. The ability to recognize words seems to develop faster than that of ascribing meanings and relations. The groups largely foreign in make-up reveal marked weakness in the latter.
3. The general intelligence of school children seems adequately measured by the Trabue Completion Test. The regularity of increase throughout the grades argues that the schools are successfully counteracting the handicap of nationality so marked in the lower grades.
4. The correct use of spoken idioms needs continued attention through the eight grades. These speech habits cannot be taken for granted.

*Arithmetical abilities.*—

1. The entire range of these special-ability tests shows that consistent improvement is being obtained from grade to grade.
2. The city-wide scores argue, after comparison with comparable conditions in other cities, that the local work is markedly satisfactory.
3. The harder arithmetical operations are executed relatively better than the easier ones, and certain of the easier operations bear comparison unfavorably.

4. The schools appear to take too much for granted in the teaching of fractions. Comparison shows that an expectation of better scores is warranted.

The statistical findings of the tests and their interpretation suggest that considerable attention may profitably be given to the actual teaching situation. The principal and teacher may profitably concern themselves with the following questions:

How does the school compare with the national and the local standards?

Which of the grades is below the standard?

Is this inferiority due to poor preparation, insufficient attention to the subject, racial handicap, poor teaching, or poor supervision?

What grades are above the standard?

Is this superiority due to superior school.and home conditions, superior native ability, superior teaching, or an unwarranted over-expenditure of time upon this particular subject?

Are the several duplicate grades producing comparable returns?

Are any of the grades showing wide group variability and thereby revealing a need of partial reclassification?

Do certain grades need to spend more time upon certain specific lines of work, and does the course of study need to emphasize these more fully?

The teacher may profitably be concerned with the following:

How does my grade compare with the national and local standards?

How shall I interpret a possible superiority—superior native ability of the pupils, better environmental conditions, better teaching, excessive amount of time spent herein, or as chance?

How shall I interpret a possible inferiority—inferior native abilities or environmental handicap, too little attention to the subject factors of chance, or poor teaching?

If the last named, how can I improve the condition—give more time to the item of demonstrated weakness, present material in a different manner, work more with lower fourth of group and give individual attention herein, and be attentive to errors of demonstrated frequency of occurrence and persistency?